LA-UR-01-3834

AN ANALYSIS OF CALIBRATION CURVE MODELS FOR SOLID-STATE HEAT-FLOW CALORIMETERS

Philip Hypes, David S. Bracken
Safeguards Science and Technology Group
Los Alamos National Laboratory, Los Alamos NM 87545

George McCabe
Department of Statistics
Purdue University, West Lafayette, IN 47907

Presented at the
Institute of Nuclear Materials Management
42nd Annual Meeting
Indian Wells, California
July 15-19, 2001

An Analysis of Calibration Curve Models for Solid-State Heat-Flow Calorimeters

Philip Hypes, David.S. Bracken
Safeguards Science and Technology Group
Los Alamos National Laboratory, Los Alamos, NM 87545

George McCabe
Department of Statistics
Purdue University, West Lafayette, IN 47907

ABSTRACT

Various calibration curve models for solid-state calorimeters are compared to determine which model best fits the calibration data. The calibration data are discussed. The criteria used to select the best model are explained. A conclusion regarding the best model for the calibration curve is presented. These results can also be used to evaluate the random and systematic error of a calorimetric measurement.

INTRODUCTION

A linear/quadratic model has been used for decades to fit the calibration curves for wheatstone bridge calorimeters. Excellent results have been obtained using this calibration curve model. The Multical software package uses this model for the calibration curve. The choice of this model is supported by 40 years [1] of calorimeter data. There is good empirical support for the linear/quadratic model. Calorimeter response is strongly linear. Calorimeter sensitivity is slightly lower at higher powers; the negative coefficient of the x^2 term accounts for this.

The solid-state calorimeter is operated using the Multical [2] software package. An investigation was undertaken to determine if the linear/quadratic model is the best model for the new sensor technology used in the solid-state calorimeter.

THE DATA

At the time this analysis was performed, 84 measurements of NIST-traceable heat standards had been made with the Solid State Calorimeter [3] developed in the Safeguards Science and Technology Group at Los Alamos National Laboratory. Most of the low-power measurements were made on single standards. Most of the high-power measurements were made by placing two or three standards in the sample cell of the calorimeter and measuring the total power produced by the standards. The power range covered by these single and combined standards was 0.72 milliwatts to 6.98 watts. Multiple measurements were made at each of the power levels. The number of measurements taken at each power level is shown in Table I. The data are plotted on a log-log plot in Fig. 1. The multiple measurements at each power level are not apparent in the plot because the variation in the signal at each power level is smaller than the plotting symbol.

Appropriate baseline values, averaging approximately 3 nanovolts, have been subtracted from the calorimeter output values, in accordance with standard calorimetry practice.

Table I. Power Levels and Replications

Power (mW)	Measurements	Power (mW)	Measurements	Power (W)	Measurements
0.72	6	78.6	5	0.81	4
1.97	7	150	5	1.25	10
1.96	5	230	5	2.5	3
4.33	5	420	8	3.74	5
17.40	5	570	5	6.98	6

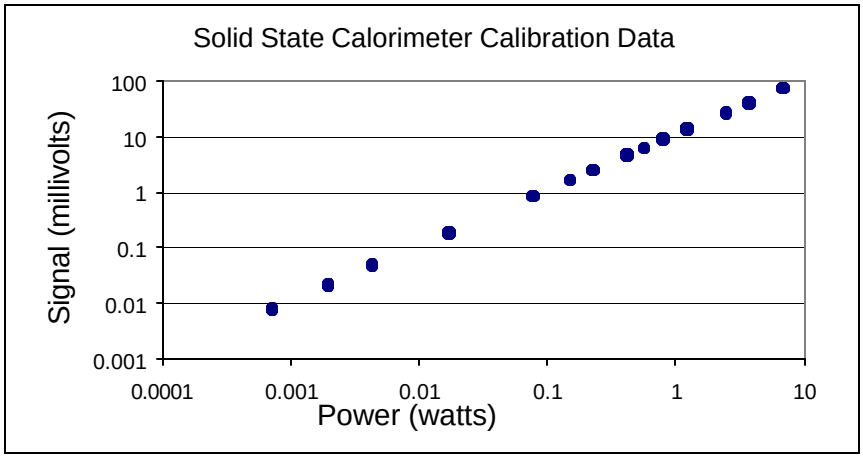

Fig. 1. Calibration data

MODELS

Regression models are generally evaluated using several accepted figures of merit. The R^2 value quantifies the degree to which the explanatory variable(s) predict the behavior of the response variables. For all models discussed below (except the first logarithmic model), the calculated R^2 values were reported as 1.000. The R^2 value was therefore of no value in choosing the best model. The F statistic indicates the strength of the relationship between the explanatory and response variables. Weak relationships result in low F values. Strong relationships result in high F values. These could be useful in discriminating between models, but only if all the models being compared have the same number of explanatory variables. Because this is not the case, we cannot use the F values to compare the models.

Residuals provide an excellent indication of the quality of the fit provided by a model. Residuals are the differences between values calculated from the fitted equation and known values. They are more subjective than the figures of merit discussed above, but they make important trends in the data very clear. They also demonstrate the agreement between each data point and the fitted line.

Standard regression analysis requires that the error distribution be constant across the range of data. The error terms should be independently and identically normally distributed, and the distribution should be centered around zero. Error terms with such a distribution would lead to residuals with a uniform distribution across the range of data. Trends in the values or variation of residuals indicate that the normal error distribution assumption is not valid.

Figure 2 provides an example of normally distributed residuals. This plot was produced using a random number generator. The random numbers represent residual values and were taken from a

standard normal distribution. The random residuals were then assigned to the power values shown in Table I.

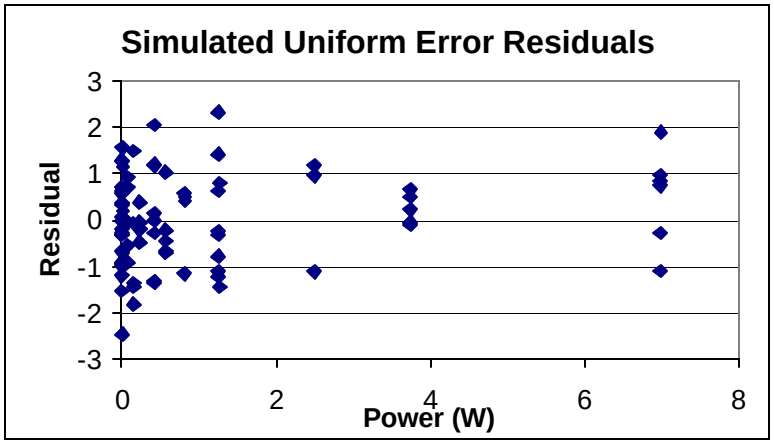

Fig. 2. An Example of Random Residuals

The highly linear nature of calorimeter response suggests a simple linear model ($y = B_1x$) as the starting point. An intercept is not included in the model because we monitor the calorimeter baseline value and subtract the baseline from each measurement result. This means that the signal (net output) for zero power will be zero voltage.

Standard regression analysis tests indicate that virtually all of the behavior of the response variable (the output voltage) is explained by the model. However, the residuals shown in Fig. 3 show that there is room for improvement. Ideally, the residuals from a fit should be uniformly distributed around zero, with no apparent structure across the range of the residuals.

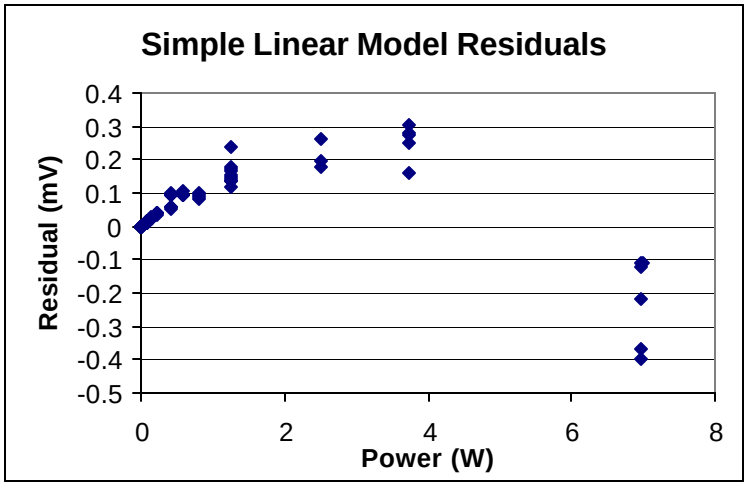

Fig. 3. Linear Model Residuals

The residuals are not constant across the power range. The pattern indicates the need for a nonlinear factor in the model. This agrees with the observation that calorimeter sensitivity decreases slightly with increasing power.

Adding a B_2x^2 term to the model should account for the nonlinearity in the data. The complete model is therefore $y = B_1x + B_2x^2$. This is the model used by the Multical software package. The residuals from this model (see Fig. 4) are clearly an improvement over the simple linear model. All the residuals are now uniformly distributed around zero. The residuals become more widely dispersed as power increases. This is due to the fact that the noise of a calorimeter measurement increases as power increases.

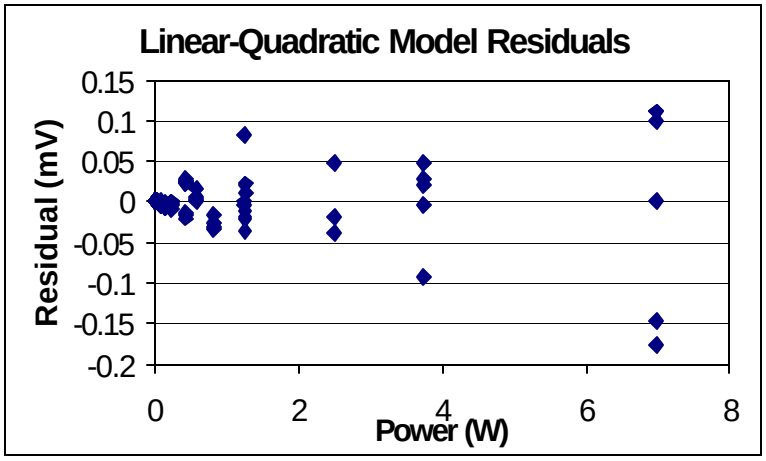

Fig. 4. Residuals from the Unweighted Linear-Quadratic Model

Standard regression methods as employed in this paper require that errors (and hence, residuals) be independently and identically normally distributed. When this requirement is violated, Bowen and Bennett [4] recommend two possible remedies (data transformation and weighted regression). Both of their recommendations were implemented.

Bowen and Bennett recommend a logarithmic transformation as being potentially useful when the error in the response variable is proportional to the value of the explanatory variable. The resulting $\ln(y) = B_0 + B_1x$ model produces residuals that are highly structured. The data presented in Fig. 5 shows that they also cover the largest range of any model included in this investigation.

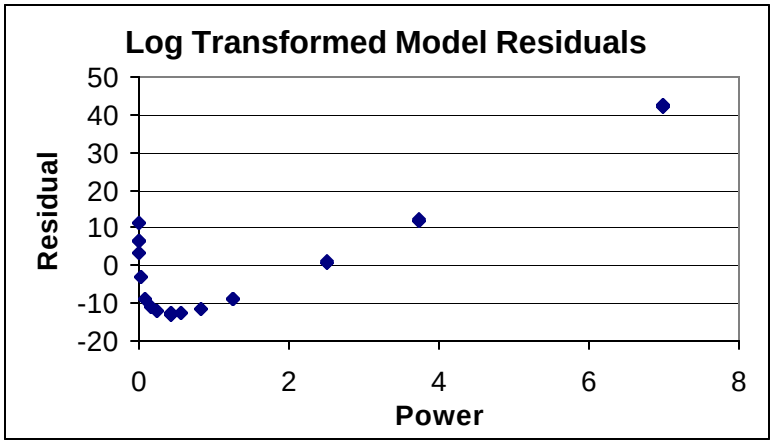

Fig. 5. Residuals from Log-Transformed Response Model

If both the response and explanatory variables are logarithmically transformed, the results improve considerably, as shown in Fig. 6. This model is $\ln(y) = B_0 + B_1 \ln(x)$.

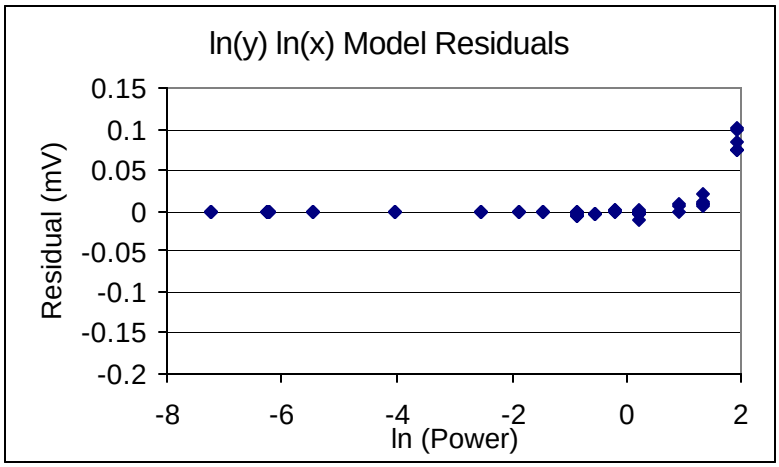

Fig. 6. Residuals from Log-Log Model

These residuals cover less of a range than in any other model. The low-power residuals are excellent. However, as the power increases, the variation in the residuals increases slightly. The positive trend over the highest power levels is also a concern. This may indicate that this model will be biased at these power levels. It may be due to the slight changes in calorimeter sensitivity for items of differing powers.

The other recommendation for dealing with non-normal residuals was weighted regression. The assumption behind weighted regression analysis is that data are more reliable where the variance is lower. Weighted regression is an iterative process in which the weight assigned to a particular data point for the final regression is inversely proportional to the square of the residual for that data point in the first regression. Those data points with small initial residuals presumably represent more reliable data and are given higher weights in the final analysis.

A weighted regression was performed using the linear/quadratic model. The weights were assigned based on the residuals already shown for the linear/quadratic model. The parameters (and therefore the residuals) obtained from the weighted regression did not differ significantly from the parameters obtained from unweighted regression. This implies that the linear/quadratic model is not affected by the increasing variability of the residuals with increasing power. This increases our confidence in the appropriateness of this model; the physical basis of the model appears to be strong enough to obtain a good fit, even when the analysis is based primarily on the lower range of the data.

We will compare the linear/quadratic model and the log-log model by examining how they perform as calibration curves for the calorimeter. Power values were calculated using each of these calibration models for a variety of voltages. The calculated power values are compared to the nominal values in Table II. In this table, model 1 is the linear/quadratic model, and model 2 is the log-log model.

Table II. Linear-Quadratic Model/Log-Log Model Prediction Comparison

Input (mV)	Calculated Power (Watts) Model 1	Calculated Power (Watts) Model 2	Nominal Power	Bias Ratios Model 1	Bias Ratios Model 2	Absolute Value of the Bias	
0.05	0.0043	0.0043	0.0043	0.9951	0.9946	0.49%	0.54%
0.18	0.0174	0.0174	0.0174	0.9993	0.9977	0.07%	0.23%
0.83	0.0784	0.0787	0.0786	1.0021	0.9993	0.21%	0.07%
1.59	0.1511	0.1516	0.1514	1.0024	0.9992	0.24%	0.08%
2.42	0.2295	0.2303	0.2298	1.0014	0.9980	0.14%	0.20%
4.41	0.4187	0.4201	0.4201	1.0034	1.0000	0.34%	0.00%
6.02	0.5725	0.5744	0.5716	0.9984	0.9952	0.16%	0.48%
8.54	0.8117	0.8141	0.8142	1.0031	1.0002	0.31%	0.02%
13.17	1.2533	1.2559	1.2466	0.9946	0.9925	0.54%	0.75%
26.16	2.4988	2.4971	2.4968	0.9992	0.9999	0.08%	0.01%
39.07	3.7444	3.7304	3.7393	0.9986	1.0024	0.14%	0.24%

The average absolute value of the bias for the linear/quadratic model was 0.246%. For the log-log model, the average absolute value of the bias was 0.239%. These values are, for practical purposes, identical.

The fitted equations from the two models were used to calculate power values for each of the output voltages observed at powers of 1.97 mW, 0.42 W, 1.25 W, and 3.74 W. The Mean Square Error of each set of observations was determined according to the formula

$$MSE = \sqrt{\frac{\sum_{i=1}^{n} \left(Y_i - Y_{avg}\right)^2}{n}}$$

where Y_i is the calculated power value for each output voltage, Y_{avg} is the average of the calculated values at each output voltage, and n is the number of calculated values at each power value. The results are shown in Fig. 7.

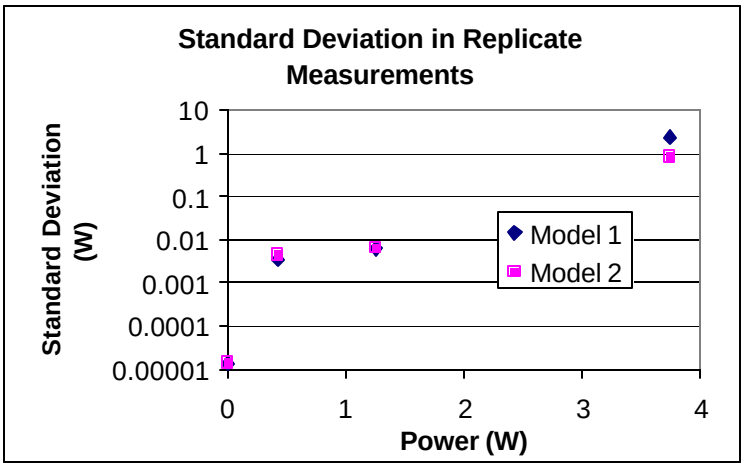

Fig. 7. MSE comparison.

The MSE values are, for practical purposes, identical for the lower power data sets. At the highest power, model two produces data with a lower MSE than model one. As is apparent in the residuals, model two produces more consistent data at the highest power level. However, while the data is more consistent, the residuals also imply that the model two results are biased, while no such bias is apparent in the model one residuals. The model one residuals display greater variation, but are centered on zero. This implies that they are unbiased.

CONCLUSION

The linear/quadratic model ($y = B_1 x + B_2 x^2$) has strong empirical support. It has been used successfully for years. A rigorous investigation confirms that it is the best model for calorimetric calibration, despite violation of the uniformity of errors requirement. It is unaffected by the lack of residual uniformity and is bias-free across the power range, unlike the log-log model. The fitted equation for the solid-state calorimeter data is

$$y = 10.54054 \pm 0.00811x - 0.02855 \pm 0.00121x^2$$

The Multical software package uses model one.

ACKNOWLEDGMENT

This work is supported by the US Department of Energy, Office of Nonproliferation and National Security, Office of Safeguards and Security.

REFERENCES

1. J. F. Eichelberger, K. C. Jordan, S. R. Orr, and J. R. Parks, *Physical Review 96,* **719** (1954).

2. R. S. Biddle and C. S. Schneider, "Multical -- Software System for Calorimeter Control," *Nucl. Mater. Manage.* **XXIX,** CD-ROM (Proc. Issue, 2000).

3. D. S. Bracken and P. Hypes, "Solid-State Calorimeter," *Nucl. Mater. Manage.* **XXIX**, CD-ROM (Proc. 2000).

4. M. Bowen, C. Bennett, “Statistical Methods for Nuclear Material Management,” Pacific Northwest National Laboratory Report No. PNL-5849 (NUREG/CR-4604), December 1988, p. 343.